

---

## **Abteilung I: Genus und Geschlecht im Lexikon**



Carolin Müller-Spitzer und Henning Loblin

# ***Leben, lieben, leiden: Geschlechterstereotype in Wörterbüchern, Einfluss der Korpusgrundlage und Abbild der sprachlichen ‚Wirklichkeit‘***

**Zusammenfassung:** Wissenschaftlich basierte allgemeine Wörterbücher des Deutschen werden heute meist korpusbasiert erarbeitet, d. h. die in ihnen beschriebene Sprache wird vor der lexikografischen Beschreibung empirisch erforscht. Diese Korpora sind allerdings, wie die großen linguistischen Textsammlungen zum Deutschen allgemein, durch Zeitungstexte dominiert. Daher beruhen die in Wörterbüchern beschriebenen Kollokationen und typischen Verwendungskontexte zumindest teilweise auf dieser Textsorte. Wir untersuchen in unserem Beitrag anhand einer Fallstudie zu *Mann* und *Frau*, wie stark sich die Beschreibung solcher Kollokationssets ändern würde, wenn als Korpusgrundlage nicht Zeitungen, sondern Publikumszeitschriften oder belletristische Texte herangezogen würden und wie unterschiedlich demnach Geschlechterstereotype dargestellt würden. Damit diskutieren wir auch die Frage, ob Zeitungstexte in diesem Fall ein adäquates und vielseitiges Abbild des Gebrauchsstandards zeigen. Auf einer allgemeineren Ebene wird dadurch ein grundlegendes Problem korpuslinguistischer Forschungsarbeiten tangiert, nämlich die Frage, inwieweit durch Korpora überhaupt ein ‚objektives‘ Bild der sprachlichen Wirklichkeit gezeichnet werden kann.

## **1 Einleitung**

Die psychologische Forschung legt nahe, dass Einordnungen von Personen schnell und abhängig von der unmittelbar verfügbaren Information gefällt werden (vgl. z. B. Baum et al. 2018). Das Erfolgskriterium dieses von Kahnemann sogenannten ‚schnellen Denkens‘ ist „die Kohärenz der Geschichte, die es erschafft. Die Menge und die Qualität der Daten, auf denen die Geschichte beruht, ist weitgehend belanglos“ (Kahnemann 2011: 112). Gleichzeitig kann dieses

---

**Anmerkung:** Wir danken unseren Kolleg\*innen Frank Michaelis, Alexander Koplenig, Sascha Wolfer, Annette Klosa-Kückelhaus, Stefan Engelberg und Petra Storjohann sowie Kathrin Kunkel-Razum für wertvolle Diskussionen zu diesem Beitrag.

Open Access. © 2022 Carolin Müller-Spitzer und Henning Loblin, publiziert von De Gruyter.

 Dieses Werk ist lizenziert unter der Creative Commons Namensnennung 4.0 International Lizenz.

<https://doi.org/10.1515/9783110746396-002>

schnelle Denken über Personen sehr stark, aber gleichzeitig unbewusst von Stereotypisierungen beeinflusst werden.

Stereotype dienen als einfache Entscheidungsregeln – Faustregeln –, die uns erlauben, Informationen leichter zu verarbeiten, aber oft treffen sie nicht zu. Schlimmer noch: Stereotype, die beschreiben, wie wir glauben, dass die Welt sei, verwandeln sich häufig in Vorschriften, wie die Welt sein sollte. Es gibt viele Belege aus der psychologischen Forschung, dass wir gar nicht anders können, als Menschen (und andere Beobachtungen) in Kategorien einzurordnen. Selten ist das ein bewusster Denkprozess, der uns Aufschlüsse über demografische Gruppen gibt. (Bohnet 2017: 15)

Stereotype verstehen wir hier allgemein als ein Denken in Gruppenkategorien, wohl wissend, dass dieses Thema in der Sozialpsychologie sehr viel ausdifferenzierter behandelt wird:

Indeed, individuals and groups can be said to be the central facts of society. Without individuals there could be no society, but unless individuals also perceive themselves to belong to groups, that is, to share characteristics, circumstances, values and beliefs with other people, then society would be without structure or order. These perceptions of groups are called stereotypes. (McGarty, Yzerbyt & Spears 2002: 1)

Solche Gruppenbeschreibungen bezüglich des Geschlechts finden sich auch in Wörterbüchern. Sehr zugespitzt und amüsant hat das Luise Pusch anhand der Beispielsätze des Duden-Bedeutungswörterbuchs aus dem Jahr 1970 gezeigt: Der Mann, also „er“, „zeigt eine akrobatische Beherrschung seines Körpers“, „seine Seele vermag das All zu umfassen“ und „große Wirkung ging von ihm aus“. „Sie“ dagegen „ist immer adrett gekleidet“, „hat das Baby täglich ausgefahren“, „erwartet mit großer Angst seine Rückkehr“ und „sie sah zu ihm auf wie zu einem Gott“. Sie resümiert dazu:

Im Vorwort schreibt die Redaktion, dass der ‚Grundwortschatz des Deutschen in seinen Grundbedeutungen‘ dargestellt werden soll. Viel mehr gelingt ihr: Sie vermittelt einen tiefen, unvergesslichen Einblick in die Seele des Deutschen, in seinen Grundempfindungs- und Grundgedankenschatz. (Pusch 1984: 144; vgl. ausführlicher zu verschiedenen Wörterbüchern zum Deutschen Nübling 2010)

An diesem Beispiel und vielen anderen zeigt sich, dass Wörterbücher oft ein Spiegel ihrer Zeit, somit auch eine der wichtigen „Bühnen für Genderinszenierungen“ (Nübling 2010: 594) darstellen. So stellen Hu, Xu & Hao (2019) in Parallele zu Nübling (2010) bei einer Analyse eines gegenwartsprachlichen chinesischen Wörterbuchs fest:

Women are often constructed in peripheral and domestic roles, as daughter, mother or grandmother. Their experiences are mostly restricted to themselves and their adjacent environment. When they act, their actions rarely bring noticeable changes to other participants.

pants or to the environment. Women are described as sensitive, loving and emotional, particularly preoccupied with familial, marital and domestic matters. On the other hand, men are mostly constructed in their central and social roles, as the prototypical adult men. [...] Men are described as strong in physical strength, versatile in skills and noble in their actions. In other words, men are represented as valuable, active social members. (Hu, Xu & Hao 2019: 28)

Unabhängig davon, ob man dies als eine adäquate Beschreibung der ‚Wirklichkeit‘ betrachtet oder als eine zu stereotype Darstellung von Männern und Frauen, stellt sich die Frage, ob solche Darstellungen von Geschlecht in Wörterbüchern gewollt sind oder sein können. So führt beispielsweise John Sinclair im Vorwort zum *Collins Cobuild English Language Dictionary* von 1987 aus, dass sie im Team „have abandoned the convention whereby *he* was held to refer to both men and women“ (Sinclair 1992: XX). Dies habe man aus verschiedenen Gründen gemacht, u. a. weil „it is a very sensitive matter for those who have pointed out the built-in sexism of English“ (Sinclair 1992: XX). Diese bewusste Positionierung sei bei Wörterbüchern insbesondere deshalb relevant, weil sie als Norminstanzen aufgefasst werden können, auch wenn sie vor allem deskriptiv intendiert sind:

This brings up the question of usage and authority. These concepts must support each other or no-one will respect either of them. If their close relationship breaks down, and authority is not backed up by usage, then no-one will respect it. [...] Similarly, no-one will respect usage if it is merely an unedited record of what people say and write. [...] Any successful record of a language such as a dictionary is itself a contribution to authority. (Sinclair 1992: XX–XXI; vgl. auch Hidalgo Tenorio 2000: 225; Barnickel 1999: 171; Ripfel 1989: 204; Kotthoff & Nübling 2018: 180)

Vor diesem Hintergrund haben Lexikografinnen und Lexikografen eine besondere Verantwortung. In der Duden-Redaktion wurde daher nach dem Erscheinen des oben zitierten Aufsatzes von Pusch versucht, an vielen Stellen das Wörterbuch zu verbessern, um möglicherweise unnötig stereotype Beispielsätze zu vermeiden und systematisch weibliche Berufsbezeichnungen aufzunehmen, wenn sie üblich sind etc. (Kunkel-Razum 2004 und 2012; Eickhoff 2012; allgemein dazu s. Westveer, Sleeman & Aboh 2018). Dabei geht es v. a. darum, eine Bewusstheit mit der Thematik zu zeigen:

Selbstverständlich sollen Wörterbücher keine asymmetrischen, im Sprachsystem verfestigten Gegebenheiten „begradi gen“. Dass beim Eintrag *Mädchen* immer auch auf das *leichte Mädchen* und beim Eintrag *Junge* auch auf den *schweren Jungen* einzugehen ist, ist unbestritten und in der deutschen Sprache (im Lexikon) verankert. Es geht nicht darum, ein *schweres Mädchen* oder einen *leichten Jungen* einzufordern [...]. Ebenso wenig geht es um schwangere Männer und weibliche Machos. Es geht hier um lexikografisches *doing gender*. [...] [D]ie Frage, welche Position auf einer Skala von *undoing gender* über

*doing gender* bis hin zu hyperritualisiertem *gender* die Wörterbücher beziehen, mit anderen Worten, welchen „Dramatisierungsgrad“ sie übernehmen – und ob sie womöglich ihrerseits eine solche Dramatisierung betreiben. (Nübling 2010: 595)

Die Darstellung von Geschlecht in Wörterbüchern scheint also in einem Spannungsfeld zwischen Sprachgebrauch und lexikografisch-moralischer Verantwortung zu liegen. Später werden wir darauf eingehen, wie sehr auch die Auswahl des Sprachgebrauchs, d. h. die Auswahl der Korpusgrundlage, einen Einfluss hat. Doch zunächst werden wir im folgenden Abschnitt einen Blick auf die aktuelle Praxis gegenwärtssprachlicher Wörterbücher werfen.

## 2 Ein Blick in die aktuelle lexikografische Praxis

Den Anfang oder Anlass dieser Untersuchung bildete ein eher abseitiges Wörterbuch: das Langenscheidt-Wörterbuch *100 % Jugendsprache* (Langenscheidt 2017), in dem klar sexistische Ausdrücke wie *Einwegtussi* oder *Standgebläse* aufgeführt werden, die – bezogen auf Mädchen bzw. Frauen – nicht als abwertend eingordnet wurden. Der einzige abwertende Ausdruck, der sich auf Jungen bzw. Männer bezieht (*Dein Penis ist nur auf Englisch dick!*), wird dagegen als „beleidigend“ gekennzeichnet. Hier kurze Auszüge aus dem Wörterbuch *100 % Jugendsprache* (Langenscheidt 2017):

- „*Einwegtussi* Frau, die ausschließlich One-Night-Stands hat
- *Fotzen glotzen* Mädchen nachschauen
- *Jungschwanz* Anrede unter Jungs, Kompliment und Respekt für die Potenz des Anderen
- *Spermienator* Mann, der viele One-Night-Stands hat
- *Ständerstimmung* gute Laune
- *Standgebläse* kleines Mädchen“

Nun ist dieses Wörterbuch wahrscheinlich keines der zentralen Produkte im Langenscheidt-Programm und sicher auch keine wichtige Ressource zum Sprachenlernen. Trotzdem ist es vorstellbar, dass Deutschlerner\*innen ein solches Buch zur Hand nehmen würden mit dem Ziel, sich besser unter Jugendlichen verständigen zu können. Da ein renommierter Verlag Herausgeber ist, könnte man gleichzeitig davon ausgehen, dass die darin verzeichnete Information verlässlich ist. Wie würde ich aber als Mutter reagieren, wenn z. B. ein Gastschüler meine Tochter als „*Standgebläse*“ bezeichnen würde (von ihm aus in dem Glauben, dies wäre ein anderer Ausdruck für ein „kleines Mädchen“)? Auf die Frage an den Verlag, ob sie die oben gezeigten Paraphrasen unter den Umständen als angemessen einstufen würden, kam die Antwort:

Bitte seien Sie versichert, dass wir bei Langenscheidt keineswegs Diskriminierungen oder verbale Entgleisungen jedweder Art gutheißen. Beim Titel „100 % Jugendsprache“ handelt es sich um eine Abbildung des Ist-Zustandes der Sprachverwendung Jugendlicher im Sinne einer (nicht wertenden) „Übersetzungshilfe“, ein Wörterbuch gewissermaßen. [...] Die Auswahl neuer Begriffe (wie auch deren „Übersetzung“) orientiert sich an aktuellen Gegebenheiten [...], ohne dass wir als Verlag Einfluss auf das Ergebnis nehmen würden. (Mail einer Mitarbeiterin im Product Management, Langenscheidt GmbH & Co. KG)

Bei allen bekannten Zweifeln, ob die Wahl des Jugendwortes des Jahres und damit auch das Wörterbuch überhaupt etwas mit diesem „Ist-Zustand“ zu tun hat (vgl. Földes 2019), ist die Argumentation interessant: Der Verlag bzw. die Mitarbeiter\*innen, die das Wörterbuch zu bearbeiten haben, übernehmen mit dieser Argumentation keine Verantwortung für die lexikografischen Inhalte, da sie nur den „Ist-Zustand abbilden“.

Dieser „Ist-Zustand“ liegt aber natürlich nicht offen zutage, sondern muss empirisch ermittelt werden. Welche Daten die Grundlage für das Jugendsprach-Wörterbuch und v. a. für die oben gezeigten Paraphrasen analysiert wurden, ist von außen nicht nachvollziehbar. Im Gegensatz dazu beruhen die wissenschaftlich basierten gegenwärtssprachlichen Wörterbücher zum Deutschen wie das *Digitale Wörterbuch der deutschen Sprache*<sup>1</sup> oder *elexiko*<sup>2</sup> auf digitalen Korpora, die die empirische Basis für die Wörterbucharbeit bilden. Bei *elexiko* ist die Korpusgrundlage zumindest partiell auch öffentlich abrufbar, d. h. es ist für Nutzer\*innen transparent, auf welchen Daten die Wörterbuchartikel beruhen (Storjohann 2005),<sup>3</sup> das gleiche gilt für das DWDS. Besonders im Fall von *elexiko* wurde dieser Prozess auch ausführlich metalexikografisch dokumentiert (u. a. Klosa 2011a). Die Duden-Wörterbücher beruhen auch auf einer eigenen Korpusgrundlage, dem „Dudenkorpus“, das aber nur verlagsintern genutzt werden kann.<sup>4</sup>

Sowohl das Dudenkorpus als auch das *elexiko*-Korpus sind von Zeitungstexten dominiert.<sup>5</sup> Dass die große Masse an schriftsprachlichen Daten, die in

---

1 Digitales Wörterbuch der deutschen Sprache: [www.dwds.de](http://www.dwds.de) (zuletzt eingesehen am 8. September 2020).

2 *elexiko* (2003 ff.), in: OWID – Online Wortschatz-Informationssystem Deutsch, hg. v. Institut für Deutsche Sprache, Mannheim, <http://www.owid.de/wb/elexiko/start.html> (zuletzt eingesehen am 8. September 2020).

3 S. die Übersicht zum *elexiko*-Korpus: [https://www.owid.de/wb/elexiko/glossar/elexiko-Korpus.html](http://www.owid.de/wb/elexiko/glossar/elexiko-Korpus.html) (zuletzt eingesehen am 8. September 2020).

4 S. z. B.: „Wie kommt ein Wort in den Duden?“: [https://www.duden.de/ueber\\_duden/wie-kommt-ein-wort-in-den-duden](https://www.duden.de/ueber_duden/wie-kommt-ein-wort-in-den-duden) (zuletzt eingesehen am 8. September 2020).

5 Die Zusammensetzung des Duden-Korpus ist nicht öffentlich einzusehen (allgemeine Informationen s. [https://www.duden.de/ueber\\_duden/Partner](https://www.duden.de/ueber_duden/Partner), zuletzt eingesehen am 8. September 2020). Die Information entstammt einer mündlichen Kommunikation mit Kathrin Kunkel-Razum, der Leiterin der Dudenredaktion.

den linguistischen Korpora zum Deutschen enthalten sind, Zeitungstexte sind, hat vor allem urheberrechtliche und prozessökonomische Gründe. Zum einen muss nur ein Vertrag mit einem Rechteinhaber geschlossen werden und trotzdem können regelmäßig neue Daten hinzukommen (im Gegensatz z. B. zu Rechteverhandlungen in der Belletristik), zum anderen müssen Verarbeitungsketten nur einmalig aufgebaut werden, die dann bei der Lieferung neuer Ausgaben automatisiert wieder durchlaufen können. Vor diesem Hintergrund ist ein „Zeitungswort“ um ein Vielfaches billiger als ein „belletristisches Wort“. Gleichzeitig gibt es auch inhaltliche Gründe, die Zeitungstexte zu einer interessanten Datengrundlage machen (vgl. Abschnitt 3).

Zunächst soll jedoch gezeigt werden, wie die Artikel zu *Mann* und *Frau* in den korpusbasierten Wörterbüchern elexiko und Duden online aussehen. Der Artikelteil der „Beispiele“ in Duden online zeigen keine Stereotypisierungen in dem Sinne, dass Frauen allein mit dem Thema Schönheit und äußere Merkmale, Männer dagegen mit Macht oder Gewalt assoziiert werden. Demgegenüber sieht man insbesondere den Beispielen zu „Frau“ eine bewusste, moderne Überarbeitung an, wohingegen bei „Mann“ feste Verbindungen stehengeblieben sind, die vielleicht heute weniger Relevanz haben („junger Mann“, „der gemeine Mann“, „der dritte Mann“). Trotzdem finden sich nur bei „Frau“ Beispiele zu Patiens-Rollen („eine Frau lieben, begehrn, heiraten“), die bei „Mann“ fehlt (vgl. Tab. 1). Außerdem ist auffällig, wie eine der Herausgeberinnen dieses Bandes angemerkt hat, dass die *Frau* mit deutlich mehr beschreibenden Adjektiven versehen ist. So kann der Eindruck entstehen, dass der *Mann* die unmarkierte, neutrale Hintergrundfolie bildet, von der sich „andere“ Geschlecht abhebt und daher zahlreicher beschreibender Attribute bedarf. Ein noch stärkeres *doing gender* (i. S. v. Kotthoff & Nübling 2018: 26–39) hin zu festgelegten Rollenklichess zeigen dagegen die Bildauswahl sowie die automatisch generierten typischen Verbindungen in den Artikeln *Mann* und *Frau* in Duden online<sup>6</sup> (vgl. Abb. 1).

Zusammen mit der Auswahl der Bilder ergeben sich, um es wertfrei auszudrücken, sehr kontrastreiche Beschreibungen. Typische Adjektive zu *Mann* sind *jung, alt, reich, stark, erwachsen, mächtig, bewaffnet* und *richtig*, die typischen für *Frau* dagegen *jung, alt, schön, groß, nackt, schwanger, gnädig* und *berufstätig*. Die Illustration zu *Mann* zeigt eine Person männlichen Geschlechts im Anzug vor digitalen Zahlenreihen, d. h. vermutlich in einem beruflichen Kontext, z. B. an der Börse. *Frau* wird dagegen illustriert durch eine Frau mit offenem, wehenden Haar inmitten der Natur in eher freizeitorientierter Kleidung. Durch die Hintertür der automatisch ermittelten Kollokationen sowie die Bildauswahl

---

<sup>6</sup> Duden Online: [www.duden.de](http://www.duden.de) (zuletzt eingesehen am 25. Oktober 2019). In der aktuellen Version in Duden online (Januar 2022) sind die Bilder bei beiden Einträgen entfernt.

**Tab. 1:** Beispiele zu *Mann* und *Frau* in Duden online.

Mann: „erwachsene Person männlichen Geschlechts“ (Beispiele Duden online)	Frau: „erwachsene Person weiblichen Geschlechts“ (Beispiele Duden online)
<ul style="list-style-type: none"> <li>– ein alter Mann</li> <li>– er ist ein ganzer Mann</li> <li>– (umgangssprachlich) typisch Mann! (das entspricht ganz der männlichen im Unterschied zur weiblichen Art; so kann auch nur ein Mann denken, handeln, fühlen)</li> <li>– sei ein Mann! (zeige dich als mutiger Mann!)</li> <li>– (umgangssprachlich; als Anrede) junger Mann, können Sie mir mal helfen?</li> <li>– (veraltet) der gemeine Mann (der Durchschnittsbürger)</li> <li>– der dritte Mann (Mitspieler) beim Skat</li> <li>– ein Mann von Geist (jemand, der Geist hat)</li> <li>– er ist für uns der geeignete, richtige Mann (Mitarbeiter)</li> <li>– ein Mann des Volkes (jemand, der mit dem Volk eng verbunden ist und in seiner übergeordneten Stellung dessen Vertrauen hat)<sup>7</sup></li> </ul>	<ul style="list-style-type: none"> <li>– eine junge, kluge, starke, reife, faszinierende, gebildete, emanzipierte, berufstätige, verheiratete, schwangere Frau</li> <li>– die Frau von heute (<i>die moderne Frau</i>)</li> <li>– die Frau seiner Träume (<i>sein weibliches Idealbild</i>)</li> <li>– eine Russin war die erste Frau im Weltall</li> <li>– sie war die Chefin und damit die wichtigste Frau in der Firma</li> <li>– sie hatte sich als einzige Frau der Freiheitsbewegung angeschlossen</li> <li>– eine Frau lieben, begehrn, heiraten</li> <li>– er hat viele Frauen (<i>Freundinnen, Geliebte</i>) gehabt</li> <li>– für die Gleichberechtigung der Frau kämpfen</li> <li>– die Rolle der Frau in der Gesellschaft hat sich gewandelt<sup>8</sup></li> </ul>

kommen also stark stereotype Beschreibungen von Geschlechterrollen ins Wörterbuch, die in den Beispielsätzen etwas weniger ins Auge fallen. Bei den automatisch ermittelten Verbindungen stellt sich die Frage, inwieweit die Zeitungslastigkeit des Korpus einen Einfluss hat, denn diese Angaben werden aus dem aktuellen (zeitungslastigen) Dudenkorpus gewonnen, wohingegen der Angabebereich der Beispiele stärker von früheren Bearbeitungen des Wörterbuchs und manueller lexikografischer Analyse geprägt sind. Aus diesem Grund lohnt sich ein Blick auf die entsprechenden Einträge in elexiko, dessen zugrunde liegenden Korpus ausschließlich aus Zeitungen besteht, und dessen wichtiges Prinzip bei der Erarbeitung, alle Angaben korpusgestützt neu zu erarbeiten:

<sup>7</sup> „Mann“. In: Duden Online: <https://www.duden.de/node/93398/revision/93434> (zuletzt eingesehen am 25. Oktober 2019).

<sup>8</sup> „Frau“. In: Duden Online: <https://www.duden.de/node/50185/revision/241603> (zuletzt eingesehen am 25. Oktober 2019).



Abb. 1: Bildauswahl sowie automatisch ermittelte Adjektiv-Verbindungen zu *Mann* und *Frau* in Duden online.

*elexiko* wird grundsätzlich **korpusgestützt** erarbeitet, d. h. alle Erkenntnisse, die in die Angabebereiche einfließen, werden aus dem zugrunde gelegten Wörterbuchkorpus gewonnen: In *elexiko* gibt es keine lexikografische Angabe, die nicht im *elexiko*-Korpus belegt werden kann, es werden keine Angaben aus anderen Wörterbüchern einfach übernommen. (Klosa 2011b: 16; vgl. auch Haß 2005a: 7; Storjohann 2005)

Auch hier führt die Auswahl der frequentesten Kollokationspartner zu stark unterschiedlichen Beschreibungen von *Mann* und *Frau* (vgl. Tab. 2 und Abb. 2). Insbesondere fällt auf, dass im Artikel *Mann* die Agens-Rolle das zweite Kollokationsset bildet („Was macht ein Mann?“), hingegen bei der Frau die Patiens-Rolle („Was geschieht mit einer Frau?“) als zweites aufgeführt wird; eine Schieflage, die schon einige Forscher\*innen als *doing gender* kritisiert haben (u. a. Nübling 2010; Hu, Xu & Hao 2019; Hidalgo Tenorio 2000). Dass dies hier im Wörterbuch so dargestellt wurde, liegt an der Frequenz der Gruppen, d. h. bei *Frau* wird in den Korpustexten des *elexiko*-Korpus die Patiens-Rolle wesentlich stärker thematisiert als die Agens-Rolle.<sup>9</sup> Bei Männern ist es umgekehrt.

<sup>9</sup> Information der beteiligten Lexikograf\*innen.

**Tab. 2:** Kollokationsset „Was wird im Zusammenhang mit *Mann* bzw. *Frau* thematisiert?“ aus elexiko.

Was wird im Zusammenhang mit <i>Mann</i> thematisiert?	Was wird im Zusammenhang mit <i>Frau</i> thematisiert?
Auto	Alter
Erektionsstörung	Beruf
Feuerwehr	Brustkrebs
Fußball	Emanzipation
Gleichberechtigung	Erwerbstätigkeit
Gleichstellung	Geburt
Handball	Kinder
	Sex
	Wechseljahre

**Frau** 🔊

Lesart: 'erwachsene weibliche Person'

zur Übersichtsseite

Bedeutungs- erläuterung	Kollo- kationen	Konstruk- tionen	Sinnverwandte Wörter	Gebrauchs- besonderheiten
----------------------------	--------------------	---------------------	-------------------------	------------------------------

**Kollokationen: Wörter im Kontext**

Wie ist eine Frau?

**Was geschieht mit einer Frau?**

- bedrohen
- belästigen
- Benachteiligung
- Diskriminierung
- ermorden
- Gewalt
- heiraten
- lieben
- misshandeln
- töten
- Unterdrückung
- vergewaltigen
- Vergewaltigung
- verletzen

**Wodurch werden Frauen gefördert?**

- Beratung
- Bildung
- Gleichbehandlung
- Gleichberechtigung
- Gleichstellung

**Mann** 🔊

Lesart: 'erwachsene männliche Person'

zur Übersichtsseite

Bedeutungs- erläuterung	Kollo- kationen	Konstruk- tionen	Sinnverwandte Wörter	Gebrauchs- besonderheiten
----------------------------	--------------------	---------------------	-------------------------	------------------------------

**Kollokationen: Wörter im Kontext**

Wie ist ein Mann?

**Was macht ein Mann?**

- dominieren
- ermorden
- erschießen
- fahren
- lieben
- schießen
- schlafen
- schnarchen
- (sich) verletzen
- sitzen
- sterben
- stürzen
- töten
- vergewaltigen

**Was geschieht mit einem Mann?**

- ermorden
- erschießen
- festnehmen
- töten
- verhaften

**Abb. 2:** Kollokationen zu *Mann* und *Frau* in elexiko (zuletzt eingesehen am 8. September 2020).

Auch das Team von elexiko sieht diese stereotype Darstellung kritisch:

Es fällt bei Frau auf, dass hier relativ häufig auf den gesellschaftlichen Status der bezeichneten Person Bezug genommen wird (z. B. *alleinerziehend, arbeitslos, berufstätig, geschieden, unverheiratet*). Solche Charakterisierungen fehlen bezogen auf Männer praktisch völlig. Dafür zeigen Adjektive wie *bewaffnet, maskiert, verdächtig, verumummt*, die nur im Wortartikel Mann erscheinen, wie sich auch hier wieder das zeitungssprachliche elexiko-Korpus, das naturgemäß relativ viel Berichterstattung über Kriminalität bzw. militärische Einsätze enthält, auswirkt. Auf der anderen Seite spiegelt es aber auch die gesellschaftliche Realität, dass Männer häufiger kriminell werden als Frauen. (Klosa & Storjohann 2011: 64)

So wirkt sich

die Korpuszusammensetzung gerade auf den Angabebereich der lexikalischen Mitspieler manchmal negativ aus, weil bestimmte Themen im (zeitungssprachlichen) Korpus so stark dominieren, dass das allgemein Erwartbare unterdrückt wird. (Klosa & Storjohann 2011: 58)

Interessant wäre deshalb zu sehen, ob andere Korpuszusammensetzungen andere Ergebnisse hervorbringen würden.

Sprachliches Handeln ist immer in einen kulturellen Kontext eingebettet. Sprache existiert nur in ihrer Verwendung und diese ist stets kulturell gerahmt; zugleich werden kulturelle Fakten, kulturelle Gewohnheiten, Konzeptualisierungen und Werte durch Sprache und in der Sprache konstruiert und sedimentiert – ja archiviert (Günthner & Linke 2007: 19).

Die empirische Wörterbuchbasis, die die Grundlage der lexikografischen Arbeit bildet, bringt diesen sprachlich-kulturellen Kontext und damit auch eine bestimmte Perspektive auf die Welt ins Wörterbuch. Zwar sind Männer tatsächlich laut Kriminalstatistik krimineller als Frauen und Frauen können von Männern vergewaltigt werden, doch ob genau diese Aspekte der in den Wörterbüchern beschriebenen ‚Wirklichkeit‘ die Hauptperspektive darstellen sollten, ist zu diskutieren. In der Fallstudie in Abschnitt 4 werden wir daher untersuchen, wie stark sich die Zeitungsperspektive von anderen Textsorten unterscheidet. Zunächst wollen wir jedoch kurz ausführen, warum Zeitungstexte als besonders geeignetes Abbild des schriftlichen ‚Gebrauchsstandards‘ gelten.

### 3 Zeitungstexte als Abbild des ‚Gebrauchsstandards‘

Die Hinwendung zur korpusbasierten Arbeit war für die Lexikografie von entscheidender Bedeutung:

Anyone who worked on the COBUILD project in the early 1980s – the dawn of corpus lexicography in English – will recall the sense of excitement that accompanied the daily revelations which the corpus delivered. The concordances gave access to the core of the language, to the typical patterning and behaviour of its most frequent words, and yielded insights that the largest bank of citations could never match. This marked the beginning of a fundamental change in the lexicographic method, and now, a generation later, the notion of embarking on a serious dictionary project without a corpus at its heart is almost unthinkable. (Rundell & Atkins 2013: 1136)

Den ‚Kern der Sprache‘ zu erfassen ist allerdings ein sehr hoher Anspruch, denn die Sprache ist in ihrer tatsächlichen Vielfalt sehr schwer zu erforschen (Arppe & Järvikivi 2007; Koplenig 2017: 327–338). Bestimmte Bereiche der mündlichen Kommunikation entziehen sich aus prinzipiellen Gründen, z. B. weil sie zu privat sind, aber auch andere Bereiche der mündlichen Kommunikation und der nicht-normgerechten Schreibproduktion stehen der Forschung nicht in großen Massen zur Verfügung, auch wenn sich die Lage auch in diesen Bereichen durch den Aufbau von Korpora gesprochener Sprache oder nicht-standardnaher Texte wie Chats verbessert hat. Die große Masse an Sprachdaten in linguistischen Korpora, z. B. dem Deutschen Referenzkorpus (Kupietz et al. 2018), bestehen aus Zeitungstexten. Prinzipiell kann ein Korpus, egal wie groß es ist, auch nie die Sprache abbilden oder repräsentativ für sie sein:

Since electronic corpora became possible, linguists have been overburdened by truisms about the relation between a corpus and a language, arguments which are as irrelevant as they are undeniably correct. Everyone seems to accept that no limits can be placed on a natural language, as to the size of its vocabulary, the range of its meaningful structures, the variety of its realisations and the evolutionary processes within it and outside it that cause it to develop continuously. Therefore no corpus, no matter how large, how carefully designed, can have exactly the same characteristics as the language itself. (Sinclair 2004; vgl. auch Leech 2007)

Trotzdem gelten Zeitungstexte neben den pragmatischen Vorteilen der ökonomischen Verarbeitungsketten auch als besonders geeignete Textsorte, um den öffentlichen Sprachgebrauch bzw. den geschriebenen Standard zu untersuchen: Sie eignen sich deshalb „sehr gut, weil sie an eine Sprachgemeinschaft insgesamt adressiert und daher stark von standardisierter Sprache geprägt sind“ (Klosa 2011b: 14). Außerdem werden Zeitungen auch hinsichtlich der Textsorte

tenbreite als vielfältig angehen: „Zusammengenommen enthalten Zeitungen also die unterschiedlichsten Texte, die amts-, gebrauchs-, alltags-, bedingt fachsprachlicher, wissenschaftlicher und journalistischer Natur sind und somit ein sehr vielseitiges Bild der Sprache bieten“ (Storjohann 2005: 63). Eisenberg stellt dementsprechend fest:

Als geschriebener Standard ist der Sprachgebrauch der überregionalen Presse anzusehen. [...] Daraus folgt, dass der geschriebene Standard im Prinzip statistisch durch eine Auswertung umfangreicher Zeitungskorpora zu ermitteln ist. (Eisenberg 2007: 217; vgl. auch Schierholz 2001: 97–98)

Einschränkend bemerkt hingegen Klosa, dass man nicht „generell davon ausgehen“ sollte, dass „Zeitungstexte die deutsche Sprache insgesamt sehr gut repräsentieren“ (Klosa 2011b: 14, Fußnote 5). Rundell und Atkins stellen explizit fest, dass für ein allgemeinsprachliches Wörterbuch eine breite Bandbreite an Texten im zugrundeliegenden Korpus vorhanden sein sollte:

Since a dictionary has to give an account of all the main meanings and uses of every headword it includes, it follows that a lexicographic corpus needs to include evidence for all these uses. This in turn argues for a corpus whose constituent texts cover the full repertoire of mainstream text-types. (Rundell & Atkins 2013: 1339)

In der folgenden Fallstudie wollen wir nun untersuchen, wie vielseitig das in der Zeitungssprache gezeichnete Bild von *Mann* und *Frau* ist, wie sehr es sich von anderen Sprachausschnitten, z. B. Texten der Belletristik oder von Publikumszeitschriften, unterscheidet und wie anders daher die im Wörterbuch gezeigte „Wirklichkeit“ aussehen würde, wenn das Korpus anders zusammengesetzt wäre.

## 4 Fallstudie: Einfluss der Korpusgrundlage auf Kollokationskandidaten zu *Mann* und *Frau*

Zwei Forschungsfragen standen am Anfang dieser Fallstudie:

- Würden sich die Inhalte der Kollokationssets zu *Mann* und *Frau* erheblich ändern, wenn die Korpusgrundlage sich ändert, d. h. nicht nur Zeitungstexte analysiert werden?
- Bieten Zeitungstexte (im Fall von *Mann* und *Frau* und auch grundsätzlich) ein vielseitiges Bild der Sprache und ein adäquates ‚Abbild des Gebrauchsstandards‘?

Im Folgenden stellen wir die Methode (4.1) für die Fallstudie vor, berichten die Ergebnisse (4.2) und diskutieren sie (4.3).

## 4.1 Methode

Zunächst wurden für die Fallstudie zusätzlich zum elexiko-Korpus zwei virtuelle Korpora<sup>10</sup> in Cosmas-II, der Korpusanalyseplattform des IDS, gebildet: zum einen ein virtuelles Korpus mit belletristischen Texten, zum anderen eines mit Publikumszeitschriften, die in den Archiven für geschriebene Sprache vorhanden waren (vgl. Tab. 3).<sup>11</sup>

Wie in Tabelle 3 zu erkennen ist, sind die drei Vergleichskorpora sehr unterschiedlich, sowohl von der Anzahl der Texte als auch von der Anzahl der Wörter. Das elexiko-Korpus ist um ein Vielfaches größer als die anderen (162mal größer als das zur Belletristik und 92-mal größer als das Zeitschriften-Korpus). Das Belletristik-Korpus ist zudem diachron ganz anders ausgelegt, da nur sehr wenige belletristische Texte in den IDS-Korpora vorhanden sind und eine Beschränkung nur auf neuere Texte eine zu kleine Textsammlung dargestellt hätte. Die Publikumszeitschriften wie *Beef!*, *Brigitte Woman*, *Chefkoch* oder *Living at Home* sind jedoch wie das elexiko-Korpus neueren Datums. Im Folgenden wird das elexiko-Korpus nur als Zeitungskorpus bezeichnet, da es ausschließlich aus Zeitungen besteht.

Um abschätzen zu können, ob sich die Kollokationssets für die Wörterbuchartikel in elexiko zu *Mann* und *Frau* ändern würden, wurden die lexikografischen Arbeitsschritte vollzogen, die für eine Erarbeitung der Kollokationen in elexiko vorgesehen waren, allerdings ohne die Kollokationen nach Lesarten zu gliedern (Klosa & Storjohann 2011; Haß 2005b).<sup>12</sup> Im ersten Schritt wurden in Cosmas-II Kookkurrenzanalysen für die Lemmata *Mann* und *Frau* auf allen drei virtuellen Korpora durchgeführt, die dann manuell ausgewertet wurden. Wie in elexiko wurden dabei die ersten 250 Kookkurrenzpartner analysiert. In der

---

**10** „Unterschiedliche sprachwissenschaftliche Fragestellungen können sich aber auf sehr unterschiedliche Sprachausschnitte beziehen – insofern ist das Deutsche Referenzkorpus als eine Art *Ur-Stichprobe* zum Gebrauch der deutschen Schriftsprache konzipiert, aus der je nach Fragestellung und zugehöriger Grundgesamtheit gezielt eine ausgewogene Stichprobe zusammengestellt werden kann. Ein solches aus Texten eines bestehenden Korpusarchivs zusammengestelltes Korpus wird auch als ein *virtuelles Korpus* bezeichnet“ ([https://de.wikipedia.org/wiki/Deutsches\\_Refenzkorpus](https://de.wikipedia.org/wiki/Deutsches_Refenzkorpus), zuletzt eingesehen am 14. Januar 2020).

**11** COSMAS-II *web* Version 2.3.3, <https://cosmas2.ids-mannheim.de/cosmas2-web/> (zuletzt eingesehen am 25. Oktober 2019).

**12** Ich (CMS) danke Petra Storjohann für die Einweisung in diese Arbeitsschritte.

**Tab. 3:** Übersicht über die drei virtuellen Korpora, die für die Fallstudie gebildet bzw. verglichen wurden.

	Belletristik	Publikums- zeitschriften	Zeitungen (elexiko-Korpus)
<b>Zusammensetzung</b>	Div. Romane 2001–2006 (Mona Misko: Kindsblut, Claudia Puhlfürst: Dunkelhaft, Heinrich G. F. Schneeweiß: Aus Sternenstaub), Siegfried Lenz (Romane, Essays und Erzählungen), Thomas Mann (Romane und Erzählungen)	art, Beifl., brand eins, Brigitte, Brigitte Woman, Chefkoch, Couch, Eltern family, Eltern, Essen und Trinken, Gala, Geo, Geo Saison, Living at Home, Nido, Neon, Psychologie Heute, Schöner Wohnen	Diverse Tages- und Wochenzeitungen (Deutschland, Österreich, Schweiz)
<b>Zeitraum</b>	1909–2011	2010–2018	1999–2012
<b>Anzahl Texte</b>	1.020	54.200	11.494.400
<b>Anzahl Dokumente</b>	209	811	2.654
<b>Anzahl Wörter</b>	16.718.000	29.348.000	2.713.864.000

Artikelarbeit wurden allerdings auch Kookkurrenzpartner in niedrigeren Rängen betrachtet, wenn das aus lexikografischer Sicht sinnvoll erschien. Die Auswertung richtete sich nach den elexiko-Kollokationssets für Gattungsprädiktoren für die Untergruppe der Individuativa,<sup>13</sup> z. B. „Wie ist ein X“, in das Adjektive eingeordnet werden können oder „Was macht man bzw. was geschieht mit X?“ bzw. „Was macht X?“ für Verben. In dieser Weise wurden die nach dem Wert der Log-Likelihood-Ratio (LLR-Wert) geordneten 250 ersten Kookkurrenzen nach der jeweiligen Liste ausgewertet und als Kandidaten für ein jeweiliges Kollokationsset eingeordnet. Die Kookkurrenzen wurden nach LLR-Wert und geordnet nach Korpus und Frageset in Tabellen übertragen.<sup>14</sup> Ein Ausschnitt aus der Gesamtauswertung ist in Tabelle 4 zu sehen.

<sup>13</sup> Vgl. den Eintrag „Gattungsprädikator“ im Glossar zu elexiko: <https://www.owid.de/wb/elexiko/glossar/Gattungsprädikator.html> (zuletzt eingesehen am 25. Oktober 2019).

<sup>14</sup> Die CSV-Tabellen können per Mail bei mueller-spitzer@ids-mannheim.de angefragt werden.

**Tab. 4:** Ausschnitt aus den Auswertungstabellen der Kookkurrenzlisten, eingeordnet nach Kollokationssets. Hier: Wie ist eine *Frau*?

Wie ist eine <i>Frau</i> ? Belletristik	LLR-Wert	Wie ist eine <i>Frau</i> ? Zeitschriften	LLR-Rang	Wie ist eine <i>Frau</i> ? Zeitungen	LLR- Rang
jung	7.453	jung	6.371	jung	7.499
alt	1.620	unabhängig	657	alt	1.326
arm	516	selbstbewusst	502	älter	482
älter	422	stark	464	schwanger	464
lieb	361	schön	466	verletzt	297
schön	317	schwanger	464	vergewaltigt	271
tot	305	erwachsen	448	berufstätig	245
fremd	280	älter	448	blond	225
hübsch	262	zierlich	373	betroffen	206
schlank	239	blond	360	nackt	203
blond	229	emanzipiert	359	interessiert	177
geschieden	213	berufstätig	353	zierlich	171
attraktiv	206	alt	350	engagiert	163
zierlich	195	jünger	345	arbeitslos	135
dunkelhaarig	187	toll	243	geschieden	130
gekleidet	140	kinderlos	242	sexuell	130
schwanger	139	kurvig	235	missbraucht	
klein	139	engagiert	234	tot	113
verheiratet	131	geschieden	210	resolut	113
jünger	130	lesbisch	204	gekleidet	105
				hübsch	100

Um diese Ergebnisse besser auf einen Blick vergleichen zu können, wurden die Listen als WordClouds visualisiert, der LLR-Wert bestimmt dabei die Größe der Wörter.<sup>15</sup>

In dieser Weise wurden Kollokationskandidaten basierend auf den drei virtuellen Korpora für die Beschreibung von Eigenschaften sowie zur Agens- wie zur Patiens-Rolle ausgewertet, d. h. Filler für folgende Kollokationssets ermittelt:

- Wie ist eine *Frau*/ein *Mann*?
- Was macht eine *Frau*/ein *Mann*?
- Was geschieht mit einer *Frau*/einem *Mann*?

<sup>15</sup> Zum Generieren der WordClouds wurde „Worditout“ verwendet: <https://worditout.com/word-cloud/create> (zuletzt eingesehen am 8. September 2020).

Da nicht alle Kookkurrenzen Kandidaten für eines der drei Fragesets sind, wurden auch nicht alle der 250 ermittelten Kookkurrenzen auf diese Sets verteilt, d. h. nur ein Teil der Gesamtlisten floss in die Fallstudie ein.

## 4.2 Ergebnisse

Was *macht* eine Frau? In den belletristischen Texten agiert sie sehr viel: sie *überlegt*, *schreit*, *unterbricht*, *bemerkt*, *antwortet*, *schluchzt*, *nimmt* sich etwas, *erwidert* oder *lacht*. In den Zeitschriften *sorgt* sie für die Zukunft *vor*, sie *lebt* und *fühlt*, *fühlt sich* zu jemandem *hingezogen*, *trägt* ein *Kopftuch*, *macht Karriere* oder *verdient Geld*. In den Zeitungen kommen Handlungen einer Frau in den Top-250-Kookkurrenzen deutlich weniger vor, sie *trägt* ein *Kopftuch*, *stirbt*, *hinterlässt* etwas, *meldet sich arbeitslos*, *hilft* und *erzählt* (vgl. Abb. 3).

Was *macht* dagegen ein Mann? In den belletristischen Texten macht er ähnliche Dinge wie die Frau, er *nickt*, *fragt*, *lächelt*, *schließt* die Augen, *starrt*, *schaut*, *fragt* oder *schüttelt* den *Kopf*. In den Zeitschriften *lebt* er, *verdient Geld*, *trägt Anzüge*, *sitzt*, *redet*, *schläft mit jemandem*, *verlässt jemanden*, *belästigt* aber auch *sexuell* oder *vergewaltigt*. In den Zeitungen ist sein Handlungsspielraum deutlich kleiner, er *bewaffnet* sich, *überfällt* etwas oder jemanden, *maskiert* sich, *tötet* oder *flüchtet* (vgl. Abb. 4). Während sich also die Kollokationssets zur Agens-Rolle von *Mann* und *Frau* in der Belletristik sehr ähnlich sind, haben sie

<p style="text-align: center;">schluchzen bemerken Tür öffnen Kopf schütteln denken schreien lachen öffnen schrägen schieben nach unterbrechen</p> <p style="text-align: center;">nehmen starrn mustern schauen sich aus bloken sagen nach lächeln erwidern räcken antworten</p>	<p style="text-align: center;">leben sich hingezogen leiden sich reizig Wollzeit arbeiten sitzen wissen neigen vorsorgen arbeiten Teilzeit arbeiten denken (Geld) verdienen (Kopftücher) tragen</p> <p style="text-align: center;">fühlen sagen wollen neigen gebrören machen lieben lieben</p>	<p style="text-align: center;">verdienen (etwas) hinterlassen leiden verlassen leben alarmieren</p> <p style="text-align: center;">(Kopftuch) tragen sterben (sich) arbeitslos melden erzählen</p>
Belletristik	Zeitschriften	elexiko-Korpus

Abb. 3: Kookkurrenzen zur Agens-Rolle: Was macht eine Frau?

<p style="text-align: center;">schauen starrn <b>Augen schließen</b> Schultern zucken grinsen sich umdrehen Achseln riechen zumachen Gesichtsausdruck Stimme flüstern <b>nicken</b> sich neigen sich schütteln lächeln fragen schnüffeln Kopf schütteln sitzten</p>	<p style="text-align: center;">sexuell belästigen Vollzeit arbeiten vergewaltigen leiden Anzüge tragen lieben sagen leben</p> <p style="text-align: center;">mit jm. schlafen fragen mögen köchen ertragen rasieren wollen verlassen arbeiten reden dominieren lernen</p>	<p style="text-align: center;">töten flüchten beobachten angreifen stören <b>(sich) bewaffnen</b> ums Leben kommen überfallen <b>(sich) maskieren</b></p>
Belletristik	Zeitschriften	Zeitungen (elexiko-Korpus)

Abb. 4: Kookkurrenzen zur Agens-Rolle: Was macht ein Mann?

<p>zuflüstern sich zuwenden <b>umbringen</b> biten kennen sich scheiden</p>	<p>verlieben unterdrücken gleichsetzen vergewaltigen <b>befragen</b> kennenlernen belästigen kennen</p>	<p>töten an Brustkrebs sterben Verletzungen erleiden überfallen werden missbrauchen <b>vergewaltigen</b> umbringen bedrohen verlassen werden erschossen werden ums Leben kommen</p>
---	---	---

Belletristik

Zeitschriften

Zeitungen (elexiko-Korpus)

Abb. 5: Kookkurrenzen zur Patiens-Rolle: Was geschieht mit einer *Frau*?

<p>anhören erkennen lassen mustern beobachten <b>auf ihn blicken</b> an ihn denken erblicken umbringen betrachten beobachten</p>	<p>verlieben lieben abhängig sein bewundern sich hingezogen fühlen <b>kennenlernen</b> attraktiv finden mit ihm wohnen mit ihm schlafen heiraten</p>	<p>verletzen niedergestochen werden sterben in Verdacht stehen <b>festnehmen</b> kennenlernen vorbestraft sein angegriffen werden</p>
--	--	---

Belletristik

Zeitschriften

Zeitungen (elexiko-Korpus)

Abb. 6: Kookkurrenzen zur Patiens-Rolle: Was geschieht mit einem *Mann*?

in den Zeitungen nichts gemeinsam. In der Belletristik handelt es sich dabei hauptsächlich um Sprechaktverben bzw. um solche, die in dieser Weise verwendet werden können in Mustern wie „„Lass uns gehen“, lächelte er“. Sie bezeichnen also eher Handlungen des Interagierens oder Kommunizierens, und dies für *Frau* und *Mann* anscheinend in eher ähnlicher Weise.

Nun zur Patiens Rolle: Was geschieht mit einer *Frau* bzw. einem *Mann* (vgl. Abb. 5 und 6)? In den belletristischen Texten *bringt* man sie *um* (es sind einige Kriminalromane im Korpus), *verehrt* sie, *wendet sich* ihnen *zu*, *flüstert* ihnen etwas *zu* oder *lässt sich scheiden*. In den Zeitschriften werden Frauen *befragt*, man *lernt sie kennen*, *verliebt sich in sie*, *heiratet sie*, *belästigt* sie aber auch, oder *unterdrückt* sie. In den Zeitungen werden Frauen *vergewaltigt*, *erleiden Verletzungen*, *sterben an Brustkrebs*, werden *getötet*, *überfallen* und *missbraucht*. Männer werden in belletristischen Texten *angeblickt*, an sie wird *gedacht*, sie werden *gemustert* oder *angehört*, *betrachtet* und *beobachtet*. In Zeitschriften *verliebt* man *sich in sie* wie in Frauen, man *lernt sie kennen*, *heiratet sie* vielleicht, *liebt* und *bewundert* sie und *wohnt* mit ihnen. In den Zeitungen werden sie *festgenommen*, *verletzt*, *angegriffen* und *niedergestochen*.

Und wie ist eine *Frau* bzw. ein *Mann* (vgl. Abb. 7 und 8)? In der Belletristik ist eine Frau *jung*, *alt*, *arm*, *hübsch*, *lieb*, *schön*, *schlank*, *fremd* oder *attraktiv*, ein Mann wird in belletristischen Texten genauso als *jung* und *hübsch* wie auch als *hochgewachsen*, *mittelgroß* oder *schmächtig*, *kräftig*, *groß* und *gutaussehend* oder *gebrochen* beschrieben. In den Zeitschriften sind Frauen *jung* und *unab-*



**Abb. 7:** Kookkurrenzen zu Eigenschaften: Wie ist eine *Frau*? (ohne *jung* und *alt* in allen drei Korpora).<sup>16</sup>



**Abb. 8:** Kookkurrenzen zu Eigenschaften: Wie ist ein *Mann*? (ohne *jung* und *alt* in allen drei Korpora).

*hängig, stark, selbstbewusst, berufstätig, schön, schwanger, engagiert und emanzipiert, Männer dagegen als jung, alt, bärtig, drahtig, reich und klein, verheiratet, fremd oder untherapiert beschrieben. In Zeitungen erscheinen Frauen auch als jung und alt, aber auch als blond, berufstätig, vergewaltigt, nackt, betroffen, engagiert, interessiert und arbeitslos. Männer dagegen als stark, richtig, verletzt, bewaffnet, unbekannt, mächtig und maskiert.*

Wie unterschiedlich stark sich die einzelnen Kollokationssets aus den verschiedenen Textsammlungen überlappen, soll abschließend an zwei Beispielen verdeutlicht werden, und zwar an den Kollokationskandidaten für „Wie ist ein *Mann*?“ und „Was macht eine *Frau*?“ (eine visuelle Darstellung der Überschneidung der Kollokationssets zur Agens-Rolle findet sich in Abb. 9).

In allen drei virtuellen Korpora werden Männer als *jung, alt, älter, klein und reich* beschrieben. Nur in Belletristik und Zeitschriften sind sie *fremd, hager, erwachsen, jünger, klug, bärfig, gebrochen*. In Zeitschriften und Zeitungen sind sie *verheiratet, nackt, mächtig, richtig, verstorben und gut aussehend*. Alle ande-

**16** *Jung* und *alt* (nur Zeitungen) heben sich vom LLR-Wert so stark ab von den anderen Kollokationskandidaten, dass in der Wortwolke die anderen Wörter nicht mehr zu erkennen gewesen wären, wenn sie in der entsprechenden Größe dargestellt worden wären (vgl. auch Tab. 4).

ren Kollokationskandidaten verteilen sich auf die einzelnen Korpora. Neben *Kopftuch tragen* ist das einzige, was eine Frau sowohl in den Zeitungen wie in den Zeitschriften macht: *leben*, *lieben* und *leiden*. Dagegen teilen Belletristik und Zeitschriften die Handlungen *sagen*, *sitzen* und *denken*, Belletristik und Zeitungen dagegen haben keine Schnittmengen in der Agens-Rolle. Ausschließlich im kleinen belletristischen Korpus kommen verba dicendi wie *fragen*, *rufen*, *antworten*, *unterbrechen*, *erwidern* oder *bemerken* vor. Das Bild, das die einzelnen Korpora in Hinsicht auf die verschiedenen Rollen und Eigenschaften von Männern und Frauen vermitteln, unterscheidet sich demnach sehr stark. Trotzdem sich dieser Vergleich nur auf die wichtigsten Kookkurrenzen bezieht, kann man an den oben gezeigten Beispielen sehen, dass die Kollokatoren im elexiko-Wortartikel zu den Fragen „Wie ist ein Mann?“ bzw. „Wie ist eine Frau?“, und vergleichbar auch bei anderen Frage-Antworten-Sets, deutlich anders aussähen, wenn ein belletristisches oder ein Zeitschriftenkorpus zugrunde gelegt worden wäre.

Dabei sollte man sich immer vergegenwärtigen, dass diese Kookkurrenzverbindungen wenig über Häufigkeiten aussagen, sondern mehr über Verbindungsstärken. Dass „Frau“ in der Belletristik so stark mit „schluchzen“ kookkuriert, bedeutet dementsprechend nicht, dass oft über schluchzende Frauen berichtet wird, sondern dass ein (vermutlich niedrigfrequentes) Wort wie „schluchzen“ signifikant affin zu „Frau“ ist, d. h. bestimmte Tätigkeiten oder Eigenschaften in den Texten stark mit Frauen bzw. Männern assoziiert sind. Genauso sagen die Kollokationen noch allein nichts darüber aus, in welchen genauen Kontexten sie vorkommen, ob sie vielleicht auch in Negationen stehen etc. *Schluchzen* könnte z. B. auch als verbum dicendum vorkommen wie in folgendem Beleg:

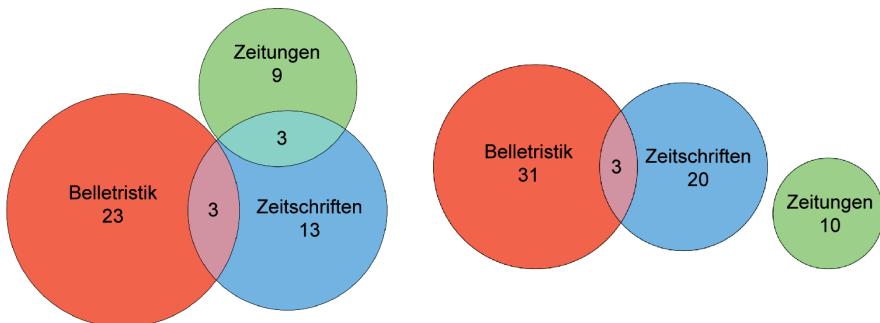
„Wir sind gar nicht verheiratet“, schluchzte die Frau plötzlich, „es hat keinen Sinn mehr, irgendwas zu leugnen, Sie finden es ja doch heraus!“ (Balàka, Bettina: Eisflüstern. – Graz, Österreich, 2006)

Diese genaue Analyse ist Teil der lexikografischen Arbeit, wie sie z. B. in einem Wörterbuch wie elexiko gemacht wird (Klosa & Storjohann 2011). Automatische Einordnungen, z. B. nach syntagmatischen Positionen, bieten auch korpusbasierte Wortprofile, wie sie z. B. das DWDS<sup>17</sup> (Didakowski & Geyken 2014) oder Korpustools wie die Sketch Engine<sup>18</sup> anbieten. In diesem Beitrag geht es v. a.

---

<sup>17</sup> <https://www.dwds.de/d/wortprofil> (zuletzt eingesehen am 10. September 2020).

<sup>18</sup> <https://www.sketchengine.eu/guide/word-sketch-collocations-and-word-combinations/> (zuletzt eingesehen am 10. September 2020).



**Abb. 9:** Überlappungen der Kollokationssets der Agens-Rolle zu Frau (links) und Mann (rechts).<sup>19</sup>

jedoch darum zu zeigen, wie groß die Unterschiede im Bereich der Kollokationen sind, wenn man die verschiedenen Subkorpora vergleicht.

### 4.3 Diskussion

Im kontrastiven Vergleich zwischen *Mann* und *Frau* wird sowohl bei der Agens- wie der Patiens-Rolle deutlich, dass die Unterschiede zwischen Mann und Frau in der Belletristik nicht besonders stark sind. Zwar sind die beschriebenen Rollen zum Teil unterschiedlich versprachlicht, aber das Handlungsspektrum deutet mehr auf einen gemeinsamen Rahmen. Bei Zeitschriften sind Unterschiede in der Patiens-Rolle weniger deutlich, in der Agens-Rolle dagegen schon klarer vorhanden, allerdings mehr auf gesellschaftliche Unterschiede bezogen, z. B. dass bei Frauen die Vorsorge ein wichtiges Thema ist. Bei beiden scheinen aber viele Handlungsverben auf, die mit Beziehungen zu tun haben. Bei Zeitungen dagegen sind die Unterschiede zwischen *Mann* und *Frau* bei beiden Rollen sehr stark und deutlich durch Gewaltkontakte bestimmt. Schaut man sich die Wortwolken zu den Kollokationskandidaten der Patiens-Rollen an, so würde bei den Belletristik-Übersichten kaum zu erkennen sein, ob es sich um Kandidaten für *Mann* oder *Frau* handelt, bei den Zeitungen ist dies dagegen auf den ersten Blick zu erkennen. Das, was Mann und Frau als Menschen gemeinsam haben, dass sie miteinander reden, streiten, sich kennenlernen oder verlieben, wird in

<sup>19</sup> Das Venn-Diagramm wurde erstellt mit dem Paket: Larsson, Johan (2019). *Area-Proportional Euler and Venn Diagrams with Ellipses*. R package version 6.0.0. <https://cran.r-project.org/package=eulerr>, zuletzt eingesehen am 25. Oktober 2019). Wir danken Sascha Wolfer für die Erstellung der Diagramme.

Zeitungen anscheinend nicht prominent thematisiert, sodass diese Handlungen in der statistischen Analyse auch nicht als relevante Kookkurrenzpartner erkannt werden.

Auch bei den Kollokationssets zu den Eigenschaften von *Mann* und *Frau* treten diese Unterschiede sehr deutlich zutage. Während in der Belletristik Frauen wie Männer in ihrem äußeren Erscheinungsbild beschrieben werden, werden diese Beschreibungen in den Zeitschriften auch aufgegriffen, aber auch gesellschaftliche Diskurse wie z. B., dass Männer *untherapiert* und Frauen *unabhängig* sind oder sein wollen. Bei Zeitungen wiederum treten ‚normale‘ oder eher alltägliche Eigenschaften eher in den Hintergrund. Hier werden mehr die Störungen des Alltags beschrieben, wie dies auch Lautenschläger beschreibt:

Deutlich wird, dass der mediale Fokus generell problemorientiert ist, dass man also, wie erwähnt, „in einer Zeitung nicht den funktionierenden Alltag findet, sondern dessen kurzzeitige Störung“ (Rüskamp 2008, 100). Daher greifen Pressetexte besonders die als problembehaftet wahrgenommenen Situation von Frauen auf, da die traditionelle Rollenverteilung mitsamt Stereotypen in beruflichen Kontexten primär Frauen zum Nachteil gereiche. (Lautenschläger 2017: 233)

Durch diese Beispiele wird deutlich, wie sehr sich die Perspektiven auf die ‚Wirklichkeit‘ in den drei Textsammlungen unterscheiden. Lexikografische Einträge zu Mann und Frau würden damit sehr unterschiedlich ausfallen, je nachdem, welche Texte als empirische Grundlage für die Wörterbucharbeit herangezogen würden.

Sprache vermittelt also nicht nur Wirklichkeit [...], sondern trägt maßgeblich zu deren Formung und Konstruktion bei und hat somit bedeutenden Einfluss darauf, wie Geschlecht ‚gemacht‘ und wahrgenommen wird. Gerade die Massenmedien sind an der Verbreitung und Verfestigung stereotyper geschlechtsspezifischer Zuschreibungen beteiligt, wobei diese nicht nur explizit, sondern auch implizit vermittelt werden. (Lautenschläger 2017: 219)

Besonders vor dem Hintergrund der eingangs dargestellten Auffassungen von stereotypen Beschreibungen und der potentiell normativen Kraft deskriptiver Wörterbücher sind die auf Zeitungstexten beruhenden Beschreibungen kritisch zu hinterfragen. Wenn die Wortartikel nicht nur so gelesen werden, wie Frauen und Männer *sind*, sondern auch, wie sie *sein sollen* oder was sie *tun sollen*, dann ist der problemorientierte Fokus der Zeitungstexte dem Anspruch einer vielseitigen Beschreibung wenig dienlich.

Außerdem treten in den Zeitungstexten die Gemeinsamkeiten von Frauen und Männern als Menschen, die viele Eigenschaften und Handlungen teilen, vor den Unterschieden zu stark zurück. Im Grunde bringt die Korpusgrundlage ein – im Vergleich zu den anderen Textsorten – unnötig starkes *doing gender* in

das Wörterbuch (vgl. dazu auch Nübling 2010: 620). Gerade für Wörterbücher ist das kritisch zu sehen:

In fact, the question is to what extent a dictionary can involve a linguistic change; or, simply, whether its role in that process must be only one of perpetuation of what is actually supported by textual evidence; in other words, why a dictionary is allowed to repeat values which imply a biased representation of reality [...]. (Hidalgo Tenorio 2000: 227)

Auch wenn man davon ausgeht, dass eine sprachliche Perspektive immer eine „biased representation of reality“ in sich birgt, hat die Fallstudie deutlich gezeigt, dass man sich mit der Auswahl der Korpusgrundlage für eine dieser sprachlichen Sichten entscheidet und dass sich diese sprachlichen Perspektiven auf die ‚Wirklichkeit‘ stark unterscheiden.<sup>20</sup> In Zeitungstexten scheinen Geschlechterstereotype besonders stark auf. Dies sind aber wiederum Differenzen, die nicht ‚an sich‘ bestehen:

Es gibt in der Realität nicht „die“ Geschlechterdifferenzen. [...] Damit geht es weder darum, reale Verhältnisse zu begradigen noch zu idealisieren noch *political correctness* zu betreiben, sondern zu bestimmten Punkten einfach keine Stellung zu beziehen – ebenso wie Wörterbücher zu (in der Realität wie in Korpora ja durchaus ebenso auffindbaren) Rassismen und Antisemiten keine Stellung beziehen (indem sie sie nicht reproduzieren). (Nübling 2010: 628)

Im Grunde ist es ein normaler Vorgang in der Wissenschaft, dass man zunächst die Chancen neuer Möglichkeiten in den Fokus rückt, z. B. dass Sprachgebrauch überhaupt intersubjektiv nachvollziehbar an großen Datenmengen untersucht werden kann. Dies war auch für Wörterbücher ein großer, wichtiger Schritt weg von der „Armchair-Linguistik“ (vgl. Fillmore 1992). Zeitungstexte schienen dafür eine besonders attraktive Textgrundlage, da sie mit überschaubarem Aufwand in großer Menge in linguistische Korpora integriert werden konnten. Ein wichtiger nächster Schritt ist es dann aber auch, die Ergebnisse dieses vielleicht zunächst etwas positivistischen Herangehens wieder zu hinterfragen. Unseres Ermessens muss noch genauer untersucht und intensiver diskutiert werden, welche Implikationen damit einhergehen, wenn z. B. ein Medium wie Zeitungen für die empirische Beschreibung des Wortschatzes dominiert.

Eine Möglichkeit könnte sein, die Methoden zur Analyse von Wortschatz für ein allgemeines Wörterbuch zumindest zu verfeinern, beispielsweise indem

---

<sup>20</sup> Annette Klosa-Kückelhaus, die sowohl beim Duden als auch im IDS als praktische Lexikografin arbeitet(e), wies bei der kritischen Lektüre dieses Beitrags darauf hin, dass sie als Lexikographin neben der Korpusgrundlage den Einfluss redaktioneller Richtlinien sowie persönlicher Bearbeitungsstrategien und -vorlieben beim Entstehen und der Redaktion einzelner Artikel nicht unterschätzen würde.

Kookkurrenzanalysen an unterschiedlichen Korpora mit unterschiedlichen Textsorten durchgeführt werden, um die daraus entstehenden Listen dann vergleichen zu können. So könnte man eher ein möglichst vielfältiges Bild erreichen, zumindest deutet unsere Fallstudie darauf hin. Da die Korpora so unterschiedlich groß sind, macht es u. E. wenig Sinn, kleine belletristische Anteile einem großen zeitungsdominierten Korpus beizumischen. Besser wäre es vermutlich, die Korpora separat statistisch zu analysieren und so die Befunde besser den unterschiedlichen Verwendungskontexten zuordnen und sie entsprechend einschätzen zu können. Dann könnte man auch genauer rückschließen, von welchen Texten welche Einflüsse ausgehen. Deshalb plädiert z. B. Sinclair auch für eine sehr feinkörnige Dokumentation aller Korpusdaten, um die Ergebnisse von Korpusanalysen besser interpretieren zu können:

Also at any time a researcher may get strange results, counter-intuitive and conflicting with established descriptions. Neither of these factors proves that there is something wrong with the corpus, because corpora are full of surprises, but they do cast doubt on the interpretation of the findings, and one of the researcher's first moves on encountering unexpected results will be to check that there is not something in the corpus architecture or the selection of texts that might account for it. (Sinclair 2004)

Die Zeitungstexte allein bieten auf jeden Fall in Bezug auf *Mann* und *Frau* kein vielgestaltiges Bild der möglichen Versprachlichung ihrer sprachlich-kulturellen Rollen. Und ob es genau dieses problemorientierte Bild ist, das die öffentliche Sprachverwendung oder die Gemeinsprache wirklich angemessen und hinreichend komplex zeigt, scheint vor dem Hintergrund der hier dargestellten Ergebnisse diskussionswürdig.

## 5 Auswirkungen auf die Korpusnutzung

Es stellt sich die Frage, wie unsere Untersuchungsergebnisse, die in den vorangegangenen Abschnitten dargestellt worden sind, für die korpuslinguistische Erkenntnisgewinnung in einem übergreifenden Sinne zu deuten sind. Gerade für das Leibniz-Institut für Deutsche Sprache (IDS) mit seinem Forschungs- und Dokumentationsauftrag bildet dies eine essentielle Frage, denn die starke Beeinflussung von Kookkurrenzdaten durch die Eigenschaften der zugrunde liegenden Korpora kann auch zu Verzerrungen in Hinsicht auf die daraus abgeleiteten Modelle führen. Es wird an dieser Stelle deutlich, dass der Ansatz, einen möglichst großen Umfang für die Korpora, auf denen eine linguistische Theoriebildung basiert, anzustreben, an seine Grenzen stößt.

Was bedeutet dies etwa für das „Deutsche Referenzkorpus“ (DeReKo) mit seinen ca. 50 Mrd. Wörtern, die vorwiegend aus Zeitungen und Zeitschriften stammen und nur zu einem vergleichsweise geringen Teil aus Belletristik oder den Sozialen Medien? DeReKo wird am IDS für verschiedene Wörterbuchprojekte eingesetzt, etwa die laufende Ermittlung von Neologismen, für die die Verfügbarkeit eines „lebenden“, ständig auf den neuesten Stand gebrachten Korpus von ausschlaggebender Bedeutung ist.<sup>21</sup> Auch verschiedene eher experimentelle lexikografische Anwendungen basieren weitgehend auf DeReKo, so dass das ‚Signal‘, das die Korpuszusammenstellung in die statistische Auswertung sendet, ganz erhebliche Auswirkungen besitzen kann (vgl. auch Kopf in diesem Band, Abschnitt 5.3.2). Der Rat für deutsche Rechtschreibung, dessen Geschäftsstelle am IDS ihren Sitz hat, greift ebenfalls auf ein aus DeReKo abgeleitetes Ratskernkorpus zurück, das zwar entsprechend der Einwohnerzahl der verschiedenen deutschsprachigen Länder und Regionen gewichtet ist, jedoch ebenfalls fast ausschließlich auf Zeitungs- und Zeitschriftentexten beruht. Das Ratskernkorpus wird gegenwärtig etwa für die Ermittlung häufiger rechtschreibschwieriger oder irregulärer Wörter genutzt oder für die Erfassung unterschiedlicher Formen gendergerechter Schreibungen.

Einen Ausweg bietet nur die Abkehr von der Vorstellung, dass ein Korpus ein homogenes, für den Sprachgebrauch repräsentatives Gebilde ist, das man ‚befragt‘ und das sodann objektive Einsichten in quantitative Strukturen einer Sprache oder in die Verwendung einzelner Lexeme bietet. Eine alternative Vorstellung beruht auf der Möglichkeit, aus einem sehr großen Korpus nur solche Teile herauszuziehen, deren Texte eine bestimmte Eigenschaft teilen, und diese wiederum zu einem eigenständigen Korpus zusammenzufassen. Im vorliegenden Beitrag ist dies für die Textgenres Belletristik, Publikumszeitschriften und Tageszeitungen aus DeReKo geschehen. Auch andere Zusammenstellungen sind denkbar, je nachdem, welche Metadaten im Gesamtkorpus erfasst sind oder welche anderen Eigenschaften schlüssig aus jedem Korpustext als Parameter abgeleitet werden können. In diesem Fall spricht man von „virtuellen Korpora“, die nicht als solche angelegt sind, sondern nach Bedarf und im Idealfall *ad hoc* gebildet werden.

An die Verwaltung von virtuellen Korpora in Textdatenbanken sind besondere Ansprüche zu stellen, die in einem ganzheitlichen Korpuskonzept nicht zum Tragen kommen. Auch auf *ad hoc* gebildete Korpora müssen die üblichen korpuslinguistischen Auswertungsverfahren effizient angewandt werden können, was hohe Anforderungen an die Indizierung des Gesamtbestandes stellt.

---

<sup>21</sup> S. <https://www.owid.de> (zuletzt eingesehen am 10. September 2020).

Am IDS wird deshalb seit einiger Zeit ein solches für die Erstellung und Auswertung virtueller Korpora optimiertes Korpusssystem entwickelt, das auf DeReKo basiert und damit alle laufenden Erweiterungen einer differenzierten Auswertung zugänglich macht. Dieses System, KorAP,<sup>22</sup> repräsentiert eine neue Generation von Korpuszugangssystemen, die den neueren Entwicklungen der korpus-linguistischen Theoriebildung, wie sie auch in dem vorliegenden Beitrag exemplarisch dargestellt worden sind, Rechnung trägt. Trotzdem bleibt die Schwierigkeit bestehen, dass Korpora unterschiedlicher Größe schlecht zu vergleichen sind (vgl. z. B. Koplenig, Wolfer & Müller-Spitzer 2019), sodass die quantitative Schieflage hin zu Zeitungen im Gesamtbestand ein nicht zu vernachlässigendes Problem darstellen kann. Damit die Konsultation unterschiedlicher virtueller Korpora für die praktische Wörterbucharbeit mit angemessenem Aufwand geschehen kann, braucht man darüber hinaus eine entsprechende Arbeitsumgebung, die ein solches Vorgehen unterstützt.

## 6 Schlussbemerkung

Im Juni 2019 schrieb die britische PR-Managerin Maria Beatrice Giovanardi einen Blogpost mit dem Titel *Have you ever googled „woman“?*, in dem sie sich vor allem über die Beschreibung von Frauen in lexikografischen Auszügen von Oxford University Press beschwerte:

The first search involved googling ‘woman synonyms’ and boom – an explosion of rampant sexism. I thought to myself, ‘What would my young niece think of herself if she read this?’ [...] Should data about how language is used control how women are defined? Or should we take a step back and, as humans, promote gender equality through the definitions of women that we choose to accept? [...] We talked about how the dictionary is the most basic foundation of language and how it influences conversations. Isn’t it dangerous for women to maintain these definitions – of women as irritants, sex objects and subordinates to men? (Giovanardi 2019)

Sie startete darauf eine Petition bei change.org, die 30.000 Unterschriften bekam (vgl. auch Vahabzadeh 2019). Interessant ist für unser Thema die Reaktion von Katherine Connor Martin von Oxford University Press, die sie gegenüber der Tageszeitung *The Guardian* folgendermaßen formulierte: Die Wörterbuchredaktion “are taking the points raised in the petition very seriously [...] As ever, our dictionaries strive to reflect, rather than dictate, language so any changes

---

<sup>22</sup> S. <https://korap.ids-mannheim.de/> (zuletzt eingesehen am 10. September 2020).

will be made on that basis” (Flood 2019). Dass die Korpusgrundlage einen erheblichen Einfluss darauf haben kann, was bei einer solchen ‚Reflexion‘ des Sprachgebrauchs herauskommen kann, hat die vorliegende Fallstudie angedeutet. Auch wenn es verführerisch ist, aus pragmatischen Gründen die am leichtesten anfallenden Daten zur Grundlage vieler wissenschaftlicher Untersuchungen und Referenzwerke wie Wörterbücher heranzuziehen, und auch wenn für viele linguistische Forschungen Zeitungstexte eine sehr sinnvolle Grundlage sein können, muss die Korpusauswahl und die Implikationen, die damit einhergehen, in unserem Fach noch intensiver diskutiert und v. a. empirisch untersucht werden.

## 7 Literatur

- Arppe, Antti & Juhani Järvikivi (2007): Take empiricism seriously! – In support of methodological diversity in linguistics [Commentary of Geoffrey Sampson 2007. Grammar without Grammaticality.]. *Corpus Linguistics and Linguistic Theory* 3(1), 99–109.
- Barnickel, Klaus-Dieter (1999): Political correctness in learners’ dictionaries. In Thomas Herbst & Kerstin Popp (Hrsg.), *The Perfect Learners’ Dictionary* (?), 161–174. Reprint 2011. Berlin, Boston: De Gruyter. doi:10.1515/9783110947021.161.
- Baum, Julia, Milena Rabovsky, Sebastian Benjamin Rose & Rasha Abdel Rahman (2018): Clear judgments based on unclear evidence: Person evaluation is strongly influenced by untrustworthy gossip. *Emotion* 20(2), 248–260. doi:10.1037/emo0000545.
- Bohnet, Iris (2017): *What works: Wie Verhaltensdesign die Gleichstellung revolutionieren kann.* (Trans.) Ursel Schäfer. 1st edn. München: C. H. Beck.
- Didakowski, Jörg & Alexander Geyken (2014): From DWDS corpora to a German word profile – methodological problems and solutions. *OPAL – Online publizierte Arbeiten zur Linguistik* 2 (2014), 39–47.
- Dudenredaktion (2019): Frau. Dudenonline. <https://www.duden.de/node/50185/revision/241603> (letzter Zugriff 25.10. 2019).
- Dudenredaktion (2019): Mann. Dudenonline. <https://www.duden.de/node/93398/revision/93434> (letzter Zugriff 25. 10. 2019).
- Dudenredaktion (2020): Wie kommt ein Wort in den Duden? Dudenonline. [https://www.duden.de/ueber\\_duden/wie-kommt-ein-wort-in-den-duden](https://www.duden.de/ueber_duden/wie-kommt-ein-wort-in-den-duden) (letzter Zugriff 08. 09. 2020).
- Eickhoff, Birgit (2012): „Frauen in den Duden“ – Werkstattbericht I aus der Dudenredaktion. In Susanne Günthner, Dagmar Hüpper & Constanze Spieß (Hrsg.), *Genderlinguistik. Sprachliche Konstruktionen von Geschlechtsidentität*, 195–212. Berlin / Boston: De Gruyter.
- Eisenberg, Peter (2007): Sprachliches Wissen im Wörterbuch der Zweifelsfälle. Über die Rekonstruktion einer Gebrauchsnorm. *Aptum* 3, 209–228.
- Fillmore, Charles J. (1992): “Corpus linguistics” or “Computer-aided armchair linguistics.” In Jan Svartvik (Hrsg.), *Directions in Corpus Linguistics*, 35–60. Berlin: De Gruyter.
- Flood, Alison (2019): Thousands demand Oxford dictionaries ‘eliminate sexist definitions’. *The Guardian*, sec. Books. <https://www.theguardian.com/books/2019/sep/17/>

- thousands-demand-oxford-dictionaries-eliminate-sexist-definitions (letzter Zugriff 13. 01. 2020).
- Földes, Csaba (2019): "Jugendwort des Jahres": kommunikative Realität oder mediale Konstruktion von Jugendsprache? *Wirkendes Wort* 69, 121–139.
- Giovanardi, Maria Beatrice (2019): Have You Ever Googled 'Woman'? *Medium*. <https://medium.com/@mbgiovanardi/have-you-ever-googled-woman-sexist-oxford-63afb87ee731> (letzter Zugriff 13. 01. 2020).
- Günthner, Susanne & Angelika Linke (2007): Linguistik und Kulturanalyse – Ansichten eines symbiotischen Verhältnisses / Linguistics and cultural analysis – aspects of a symbiotic relationship. *Zeitschrift für Germanistische Linguistik* 34 (1–2), 1–27. doi:10.1515/ZGL.2006.002.
- Haß, Ulrike (2005a): elexiko – Das Projekt. In Ulrike Haß (Hrsg.), *Grundfragen der elektronischen Lexikografie. elexiko – das Online-Informationssystem zum deutschen Wortschatz*, 1–18. Berlin: De Gruyter.
- Haß, Ulrike (2005b): Semantische Umgebung und Mitspieler. In Ulrike Haß (Hrsg.), *Grundfragen der elektronischen Lexikografie. elexiko – das Online-Informationssystem zum deutschen Wortschatz*, 227–234. Berlin: De Gruyter.
- Hidalgo Tenorio, Encarnación (2000): Gender, Sex and Stereotyping in the Collins COBUILD English Language Dictionary. *Australian Journal of Linguistics* 20(2), 211–230.
- Hu, Huiyan, Hai Xu & Junjie Hao (2019): An SFL approach to gender ideology in the sentence examples in the Contemporary Chinese Dictionary. *Lingua* 220, 17–30. doi:10.1016/j.lingua.2018.12.004.
- Kahnemann, Daniel (2011): *Schnelles Denken, langsames Denken*. 11th edn. München: Siedler.
- Klosa, Annette (Hrsg.) (2011a): *elexiko. Erfahrungsberichte aus der lexikografischen Praxis eines Internetwörterbuchs*. Tübingen: Narr Francke Attempto. <https://ids-pub.bsz-bw.de/frontdoor/index/index/docId/5154> (letzter Zugriff 13. 01. 2020).
- Klosa, Annette (2011b): Einleitung. In Annette Klosa (Hrsg.), *elexiko. Erfahrungsberichte aus der lexikografischen Praxis eines Internetwörterbuchs*, 10–26. Tübingen: Narr Francke Attempto. <https://ids-pub.bsz-bw.de/frontdoor/index/index/docId/5154> (letzter Zugriff 13. 01. 2020).
- Klosa, Annette & Petra Storjohann (2011): Neue Überlegungen und Erfahrungen zu den lexikalischen Mitspielern. In Annette Klosa (Hrsg.), *elexiko. Erfahrungsberichte aus der lexikografischen Praxis eines Internetwörterbuchs*, 49–80. Tübingen: Narr Francke Attempto. <https://ids-pub.bsz-bw.de/frontdoor/index/index/docId/5154> (letzter Zugriff 13. 01. 2020).
- Koplenig, Alexander (2017): Against statistical significance testing in corpus linguistics. *Corpus Linguistics and Linguistic Theory* 15(2), 321–346. doi:<https://doi.org/10.1515/cllt-2016-0036>.
- Koplenig, Alexander, Sascha Wolfer & Carolin Müller-Spitzer (2019): Studying Lexical Dynamics and Language Change via Generalized Entropies: The Problem of Sample Size. *Entropy* 21(5), 464. doi:10.3390/e21050464.
- Kotthoff, Helga & Damaris Nübling (2018): *Genderlinguistik: Eine Einführung in Sprache, Gespräch und Geschlecht* (Narr Studienbücher). Tübingen: Narr Francke Attempto.
- Kunkel-Razum, Kathrin (2004): Die Frauen und der Duden – der Duden und die Frauen. In Karin M. Eichhoff-Cyrus (Hrsg.), *Adam, Eva und die Sprache: Beiträge zur Geschlechterforschung*, 308–315. Mannheim: Dudenverlag.

- Kunkel-Razum, Kathrin (2012): „Er sah zu ihr auf wie zu einer Göttin“ statt „Sie sah zu ihm auf wie zu einem Gott“ (Luise Pusch)? – Werkstattbericht II aus der Dudenredaktion. In Susanne Günthner, Dagmar Hüpper & Constanze Spieß (Hrsg.), *Genderlinguistik. Sprachliche Konstruktionen von Geschlechtsidentität*, 213–220. Berlin / Boston: De Gruyter.
- Kupietz, Marc, Harald Lüngen, Paweł Kamocki & Andreas Witt (2018): The German Reference Corpus DeReKo: New Developments – New Opportunities. In Nicoletta Calzolari, Khalid Choukri, Christopher Cieri, Thierry Declerck, Sara Goggi, Koiti Hasida, Hitoshi Isahara, et al. (Hrsg.), *Proceedings of the Eleventh International Conference on Language Resources and Evaluation (LREC 2018)* 8, 4353–4360. Miyazaki: European Language Resources Association (ELRA).
- Langenscheidt, Redaktion (2017): *Langenscheidt 100 Prozent Jugendsprache 2018 – Das Buch zum Jugendwort des Jahres*. München, Wien: Langenscheidt.
- Lautenschläger, Sina (2017): (Stereotype) Männlichkeit und Weiblichkeit im Pressetext. *Osnabrücker Beiträge zur Sprachtheorie* 91, 217–236.
- Leech, Geoffrey (2007): New resources, or just better old ones? The Holy Grail of representativeness. In M. Hundt, N. Nesselhauf & C. Biewer (Hrsg.), *Corpus Linguistics and the Web*, 133–149. Leiden: Brill | Rodopi.
- McGarty, Craig, Vincent Y. Yzerbyt & Russell Spears (2002): Social, cultural and cognitive factors in stereotype formation. In Craig McGarty, Vincent Y. Yzerbyt & Russell Spears (Hrsg.), *Stereotypes as Explanations: The formation of meaningful beliefs about social groups*. Cambridge: Cambridge University Press. doi:10.1017/CBO9780511489877.
- Nübling, Damaris (2010): Zur lexikografischen Inszenierung von Geschlecht. Ein Streifzug durch die Einträge von Frau und Mann in neueren Wörterbüchern. *Zeitschrift für Germanistische Linguistik* 37(3), 593–633. doi:10.1515/ZGL.2009.037.
- Pusch, Luise F. (1984): „Sie sah zu ihm auf wie zu einem Gott“. Das Duden-Bedeutungswörterbuch als Trivialroman. *Das Deutsche als MännerSprache: Aufsätze und Glossen zur feministischen Linguistik*, 135–144. Frankfurt a. M.: Suhrkamp.
- Ripfel, M. (1989): Die normative Wirkung deskriptiver Wörterbücher. In Hugo Steger & H. E. Wiegand (Hrsg.), *Wörterbücher – Dictionaries – Dictionnaires. Ein Internationales Handbuch zur Lexikographie* (Handbücher Zur Sprach- Und Kommunikationswissenschaft), vol. 5.1, 189–207. Berlin, New York: De Gruyter.
- Rundell, Michael & B. T. Sue Atkins (2013): Criteria for the design of corpora for monolingual lexicography. In Rufus Hjalmar Gouws, Ulrich Heid, Wolfgang Schweickard & Herbert Ernst Wiegand (Hrsg.), *Dictionaries. An International Encyclopedia of Lexicography: Supplementary Volume: Recent Developments with Focus on Electronic and Computational Lexicography*, 1336–1343. Berlin, Boston: De Gruyter.
- Rüskamp, Wulf (2008): Journalistisches Schreiben. In Thomas Hauser (Hrsg.), *Zeitung machen – Zeitung lesen. Journalismus und Didaktik im Gespräch*, 98–102 Freiburg u. a.: Rombach.
- Schierholz, Stefan J. (2001): *Präpositionalattribute. Syntaktische und semantische Analysen*. Tübingen: Niemeyer.
- Sinclair, John (1992): Introduction. *Collins Cobuild English Language Dictionary*, XV–XXI. 4th edn. London: HarperCollins.
- Sinclair, John (2004): *Developing Linguistic Corpora: a Guide to Good Practice*. <http://users.ox.ac.uk/~martinw/dlc/chapter1.htm> (letzter Zugriff 13. 01. 2020).
- Storjohann, Petra (2005): Das elexiko-Korpus. Aufbau und Zusammensetzung. In Ulrike Haß (Hrsg.), *Grundfragen der elektronischen Lexikografie. elexiko – das Online-Informationssystem zum deutschen Wortschatz*, 55–70. Berlin: De Gruyter.

- Vahabzadeh, Susan (2019): Definiere: Frau. *Süddeutsche Zeitung*. <https://www.sueddeutsche.de/kultur/geschlechtergerechtigkeit-definiere-frau-1.4605718> (letzter Zugriff 13. 01. 2020).
- Westveer, Thom, Petra Sleeman & Enoch O Aboh (2018): Discriminating Dictionaries? Feminine Forms of Profession Nouns in Dictionaries of French and German\* | International Journal of Lexicography | Oxford Academic. *International Journal of Lexicography* 31(4), 371–393. doi:10.1093/ijl/ecy013.
- Wikipedia: Deutsches Referenzkorpus. In Wikipedia – die freie Enzyklopädie. [https://de.wikipedia.org/wiki/Deutsches\\_Refenenzkorpus](https://de.wikipedia.org/wiki/Deutsches_Refenenzkorpus) (letzter Zugriff 14. 01. 2020).

